



UNITED STATES ENVIRONMENTAL PROTECTION AGENCY  
WASHINGTON, D.C. 20460

APR 10 2013

OFFICE OF CONGRESSIONAL  
AND INTERGOVERNMENTAL RELATIONS

The Honorable Lamar Smith  
Chairman  
Committee on Science, Space, and Technology  
U.S. House of Representatives  
Washington, D.C. 20515

Dear Chairman Smith:

Thank you for your March 4, 2013, letter to the U.S. Environmental Protection Agency. Your letter raises concerns about access to data used by research institutions to conduct certain epidemiological studies that examine the health risks associated with exposure to fine particles and ozone pollution and requests that the EPA provide you with these data. Your letter also raises concerns regarding the Integrated Science Assessment for Ozone, which was finalized in February 2013. The enclosure provides detailed information in response to your inquiry, which I summarize below.

The EPA is committed to compliance with the requirements of the Shelby Amendment (Public Law 105-277) and to transparency with regard to the scientific bases of agency decision making, and to increasing access to federally-sponsored scientific research as outlined in the recent memorandum<sup>1</sup> from the Office of Science and Technology Policy (OSTP) in the Executive Office of the President. Both the Shelby Amendment and the OSTP memorandum recognize that increasing access to federally funded scientific data must be balanced with requirements to protect the research subject's privacy.

The EPA is transparent with regard to the scientific bases of agency decision making and disagrees with your assertion that the agency relies on "secret" data in regulatory actions and assessments of health benefits. In setting the National Ambient Air Quality Standards (NAAQS) and in assessing health benefits anticipated from air pollution regulations, the EPA relies on the scientific studies that are published in the peer-reviewed literature. The EPA provides the information used in regulatory decisions, including the epidemiological studies, in the publicly available docket accompanying each rulemaking.

Your March 4 letter requests underlying data used by researchers to conduct peer-reviewed studies of two epidemiological cohorts: the Harvard Six Cities cohort (datasets housed at

<sup>1</sup> Holdren, John P. *Memorandum for the Heads of Executive Departments and Agencies: Increasing Access to the Results of Federally Funded Scientific Research*. Office of Science and Technology Policy, Executive Office of the President. February 22, 2013.

[http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp\\_public\\_access\\_memo\\_2013.pdf](http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf)

Harvard University) and the American Cancer Society cohort (datasets housed at New York University). It is important to understand that the underlying data you are requesting for each epidemiological study consist of three distinct datasets, which the researchers link together in order to estimate the relative risks of exposure to air pollution: (1) air quality data; (2) health event data, which in these studies are data from the National Death Index; and (3) individual health data that are gathered through questionnaires completed for each study participant in the cohort. The questionnaires for these studies requested very detailed personal information, including questions on residential location, age, race, educational attainment, body mass index, alcohol consumption, smoking history, occupational exposure to pollution, and medical history. The complete, linked set of data underlying these studies is held by the scientific researchers that conducted the relevant research, not the EPA. As explained in greater detail in the enclosure, the availability of some of these datasets is subject to certain protections against disclosure of medical or similar information that could be used to identify a particular person in a research study.

As your March 4 letter notes, the EPA has previously received and responded to a similar Congressional request related to these epidemiological studies. In response to a September 2011 letter from Representative Harris, the EPA sent letters to Harvard University and New York University asking them to provide the "research data" that was funded with EPA grant dollars. As explained in the enclosure, under governing law and regulations, "research data" that the researchers are required to provide, and that the EPA is authorized to receive and disclose, excludes certain information that could be used to identify a particular person in a research study. The relevant studies are large, epidemiological research projects that received funding from a number of different sources, including the EPA, other federal agencies, and non-federal sources. Harvard University and New York University both responded to EPA's request by providing the research data that was funded with EPA grant dollars. Harvard University provided air quality data and health event data from the National Death Index. New York University provided air quality data. In June 2012, EPA sent all of the data provided by Harvard University and New York University to Representative Harris. For your reference, I have enclosed the June 7, 2012, letter from the EPA to Representative Harris, along with the data from Harvard University and New York University that has been provided.

The EPA recognizes that the data provided in response to the request from Representative Harris are not sufficient in themselves to replicate the analyses in the epidemiological studies for two reasons. First, these cohort analyses on premature mortality relied on linking private medical and demographic information with air quality data. The combination of these data could identify specific individuals and thus could not be released in its original format without consent of the study participants. Second, as noted above, it appears that some of the underlying data used in the studies that are the subject of this letter were originally collected using funds awarded to other government entities. For these reasons, composing a data set sufficient to even generally replicate the published analyses to which you refer is a complicated undertaking requiring the input of several funding agencies, awardees, and the resources of federal (or other) non-disclosure boards to ensure that the data cannot be used to identify an individual in a research study. We would welcome an opportunity to meet with your staff to discuss these matters in greater detail if you wish.

In your March 4 letter, you also requested data for several newer studies on fine particles and ozone, which are also held by research institutions and include private medical information. As outlined in the enclosure, the EPA is willing to submit a new request to the research institutions for the research data corresponding to the additional years of follow-up in the newer studies cited in your letter of March 4, 2013. We note that, because of the limitations discussed above, related to both private information and funding, the information that could be disclosed for both the older and the newer epidemiological studies would be insufficient to replicate the analyses in those studies. In addition, the enclosure describes a potential alternative approach, which has been used in the past, through which independent researchers have contacted the original researchers, entered confidentiality agreements and gained access to raw data from these epidemiological studies for purposes of undertaking reanalysis.

Finally, your letter expresses concerns regarding certain studies and causality determinations in the Integrated Science Assessment for Ozone (ISA). The ISA, which is part of the periodic review of the NAAQS that is required by the Clean Air Act, relies on a framework that has been reviewed by the public and endorsed by the Clean Air Scientific Advisory Committee. These concerns are addressed in the enclosure, which provides further information about the studies you cite, the way in which they have been evaluated under the relevant framework, and the basis for the ISA determinations.

Again, thank you for your letter. If you have further questions, please contact me or your staff may call Cheryl Mackay in my office at (202) 564-2023.

Sincerely,



Arvin Ganesan  
Associate Administrator